

Sistema software para segmentar la pizarra y el profesor en secuencias de vídeos de clases magistrales

Universidad Rey Juan Carlos I
Trabajo de fin de grado

Lara Gutiérrez Haro



ÍNDICE

1. ANTECEDENTES
2. OBJETIVOS
3. ANÁLISIS
4. DESCRIPCIÓN DEL ALGORITMO
5. PRUEBAS
6. MÉTRICAS
7. CONCLUSIONES

1. ANTECEDENTES

Este proyecto nace de la idea de mejorar la calidad con la que los alumnos atienden una clase online.

Este tipo de clases online presentan algunas dificultades:

- La **pizarra** se encuentra **alejada** de la pantalla y con **perspectiva** dificultando su lectura.
- Aparecen **elementos innecesarios** como pueden ser sillas, mesas u otros objetos.
- El **profesor se mueve** durante la clase.



¿Sería
posible
mejorarlo?

2. OBJETIVOS

Partiendo de una secuencia de vídeo 4k del frontal de una clase, generar **dos secuencias** adicionales: una que sólo contenga la pizarra y otra que sólo contenga al profesor. Este objetivo se puede desglosar en los siguientes objetivos específicos:

- **Detectar** la presencia del **profesor** y **segmentar** su imagen del resto de la escena.
- **Segmentar** la imagen de la **pizarra**, **corregir** su **perspectiva** y mejorar su relación de **contraste**.



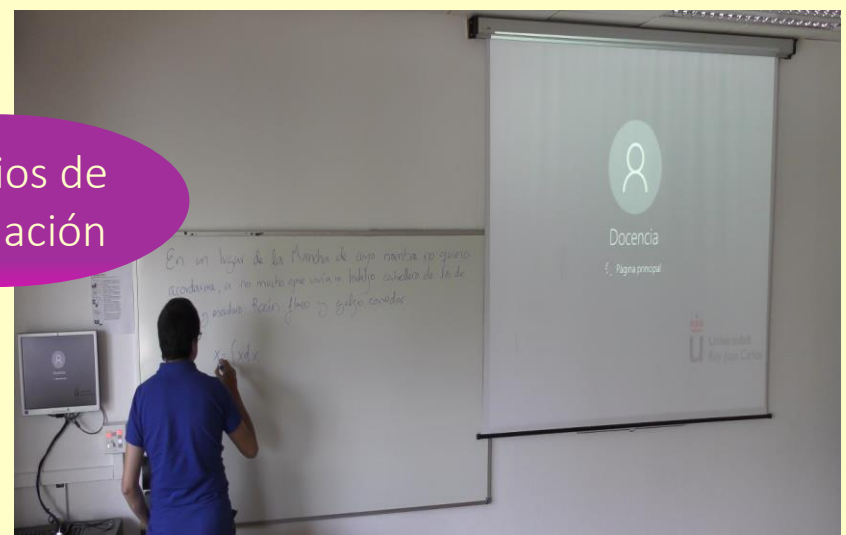
3. ANÁLISIS

Se desea construir un software:

- Para **Segmentar** la imagen del **profesor** en una secuencia de vídeo.
- Para **Segmentar** y corregir la **perspectiva** de la **pizarra**.
- Que mejore la relación de **contraste** de las **pizarras** blancas.
- Que permita **configurar** la posición de la pizarra para cada aula.
- Que sea robusto frente a los cambios de **iluminación** en el aula.
- Que procese un vídeo en tiempo cercano a **tiempo real**.
- Que utilice recursos de **hardware modestos**.
- Que funcione con vídeos de calidad **4k**.
- Que emplee bibliotecas de **código abierto**.

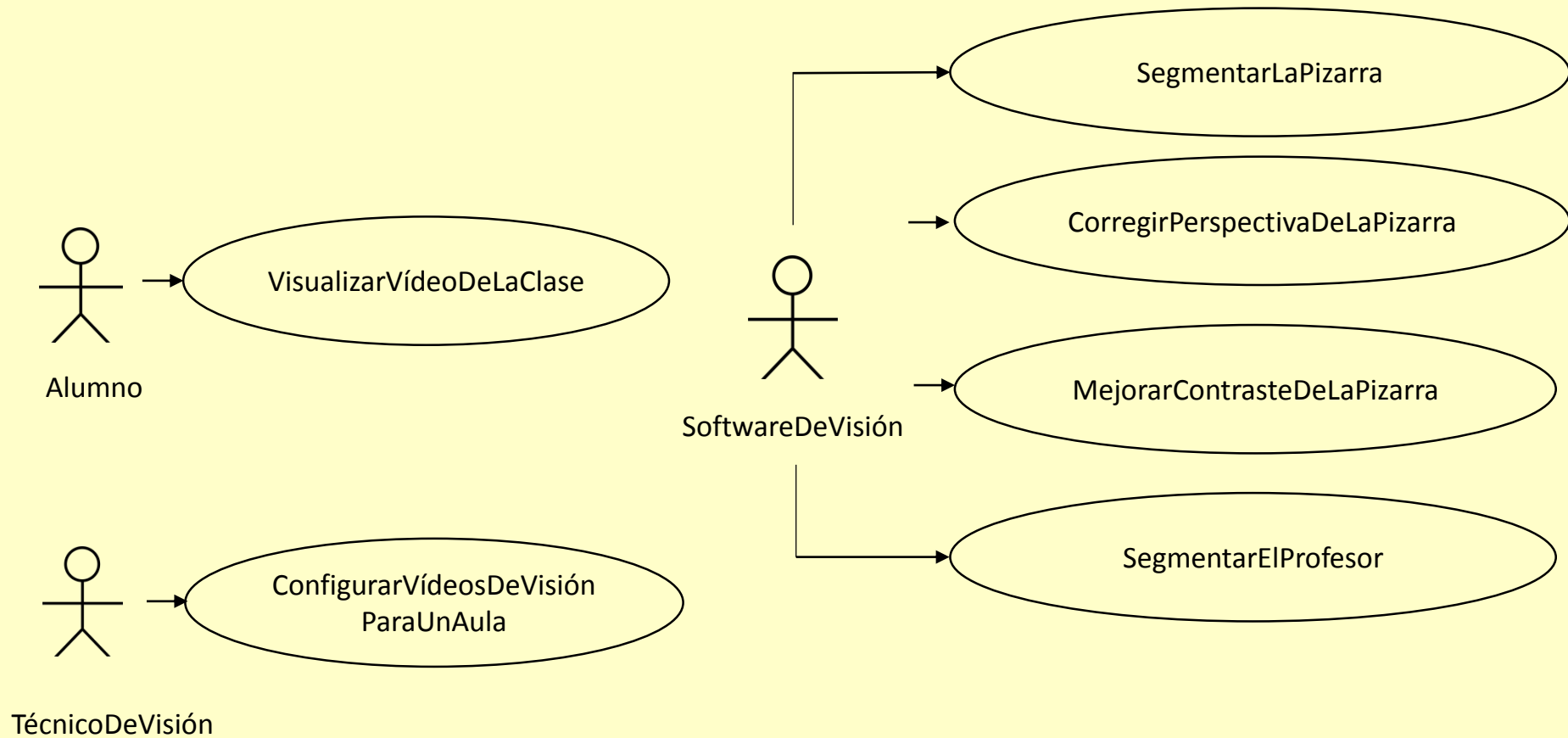


Cambios de
iluminación



3. ANÁLISIS

Diagramas de casos de Uso:



4. DESCRIPCIÓN DEL ALGORITMO

El programa consta de **dos partes** importantes respecto al software:

- Un **detector** de pantalla.
- **Seguimiento** del profesor.

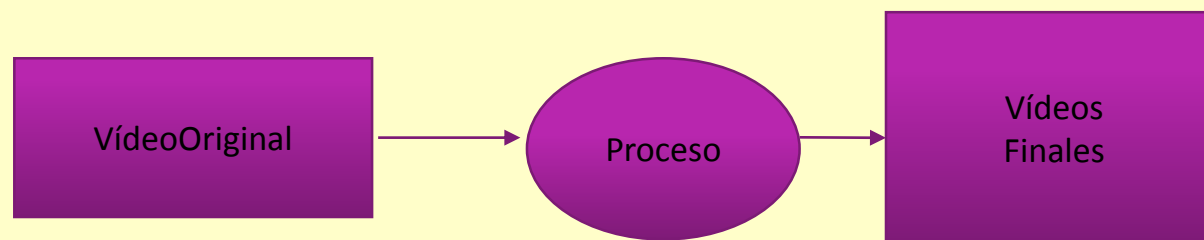
Entrada:

- Un **vídeo** de una clase grabado anteriormente.

Salida:

- Vídeo de la **pizarra**.
- Vídeo del **profesor**.

Respecto al hardware el único elemento necesario es un ordenador por lo que no supone nuevos costes a la universidad.



4. DESCRIPCIÓN DEL ALGORITMO

Segmentación de la pantalla:



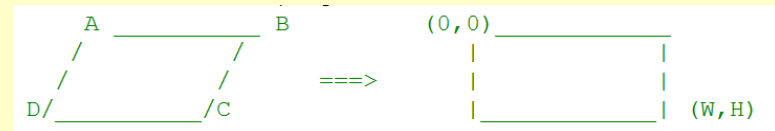
Entrada: vídeo
de una clase

Búsqueda de
extremos en
cada frame

Transformación
Afín

Mejora de la
calidad

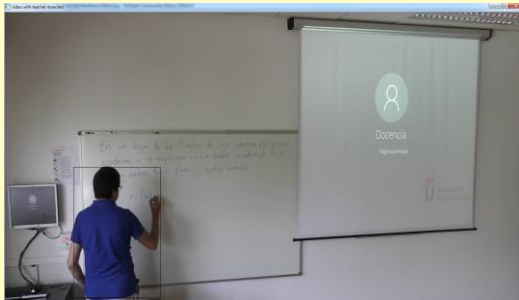
Salida:
La pizarra



```
<xml>
  <PizarraImagen a_x="172" a_y="78" b_x="689" b_y="81" c_x="672" c_y="354" d_x="237" d_y="353"/>
  <PizarraReal width="100" height="50"/>
</xml>
```


4. DESCRIPCIÓN DEL ALGORITMO

Segmentación del profesor:



Entrada:
Vídeo de
un aula

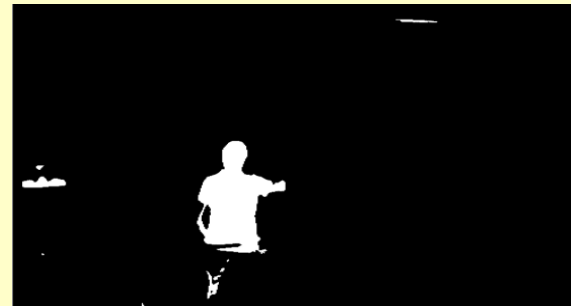


Sustracción
de fondo

Filtrado de
imagen

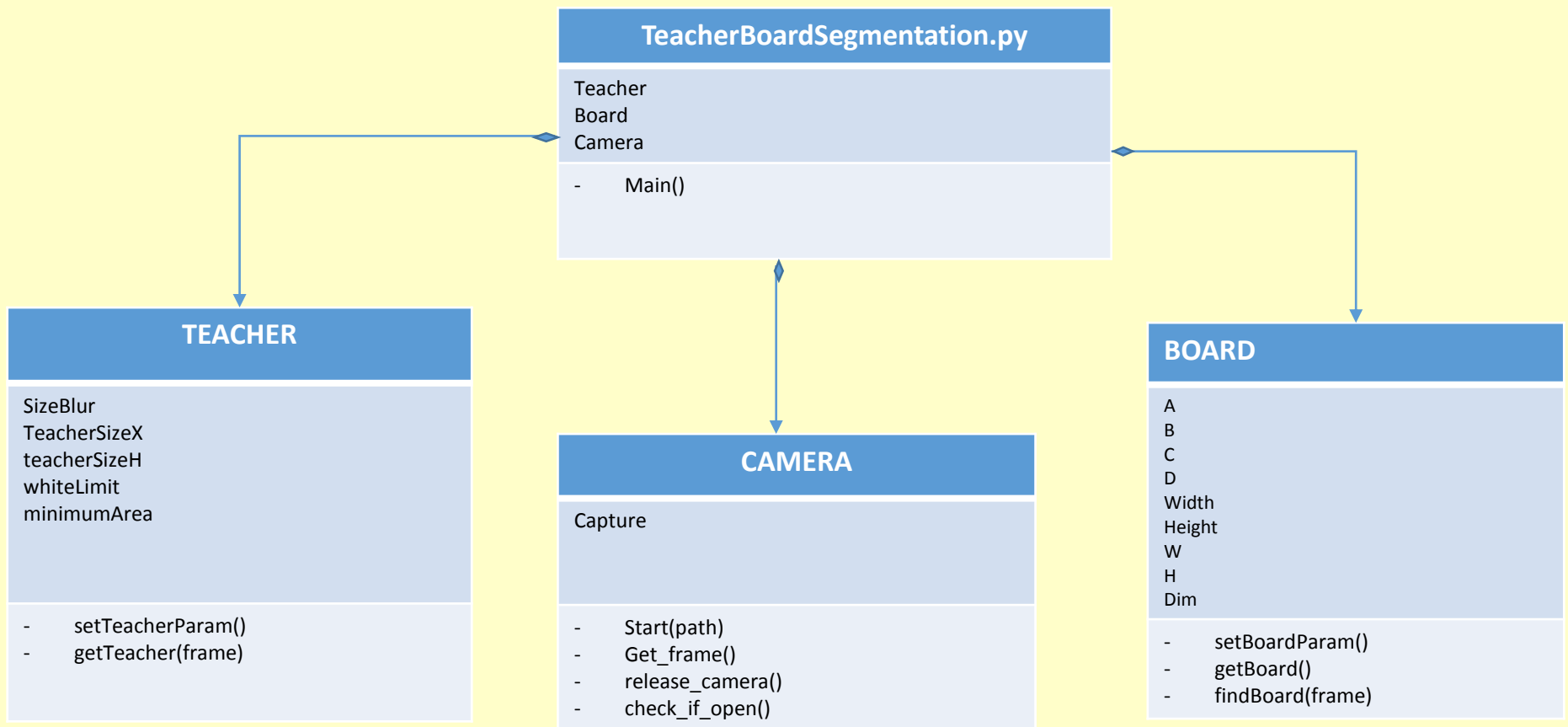
Búsqueda de
contorno

Filtro de
contorno



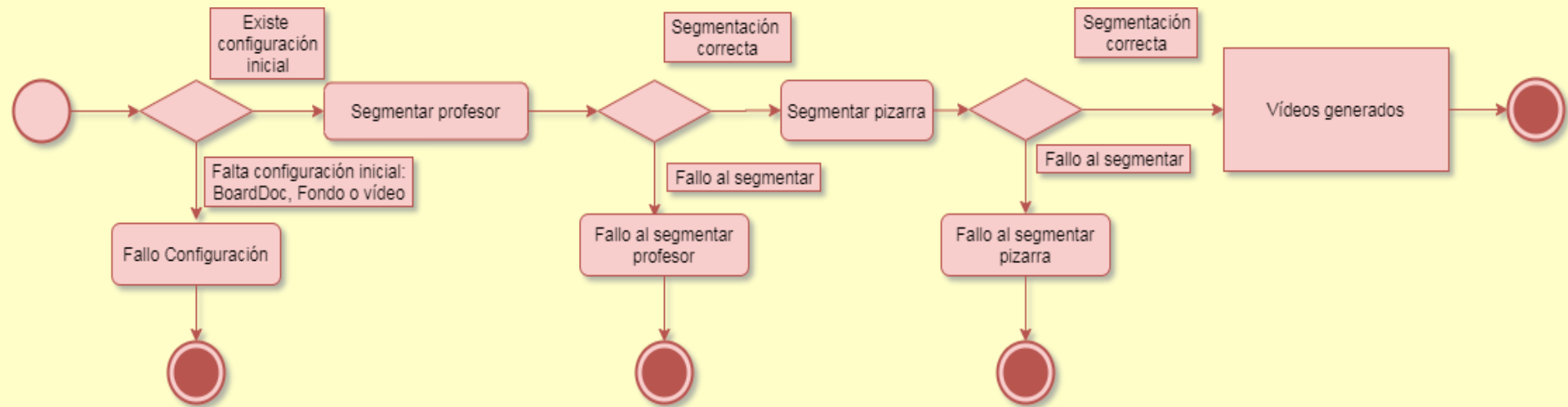
4. DESCRIPCIÓN DEL ALGORITMO

Diagrama de clases:



4. DESCRIPCIÓN DEL ALGORITMO

Diagrama de actividad:

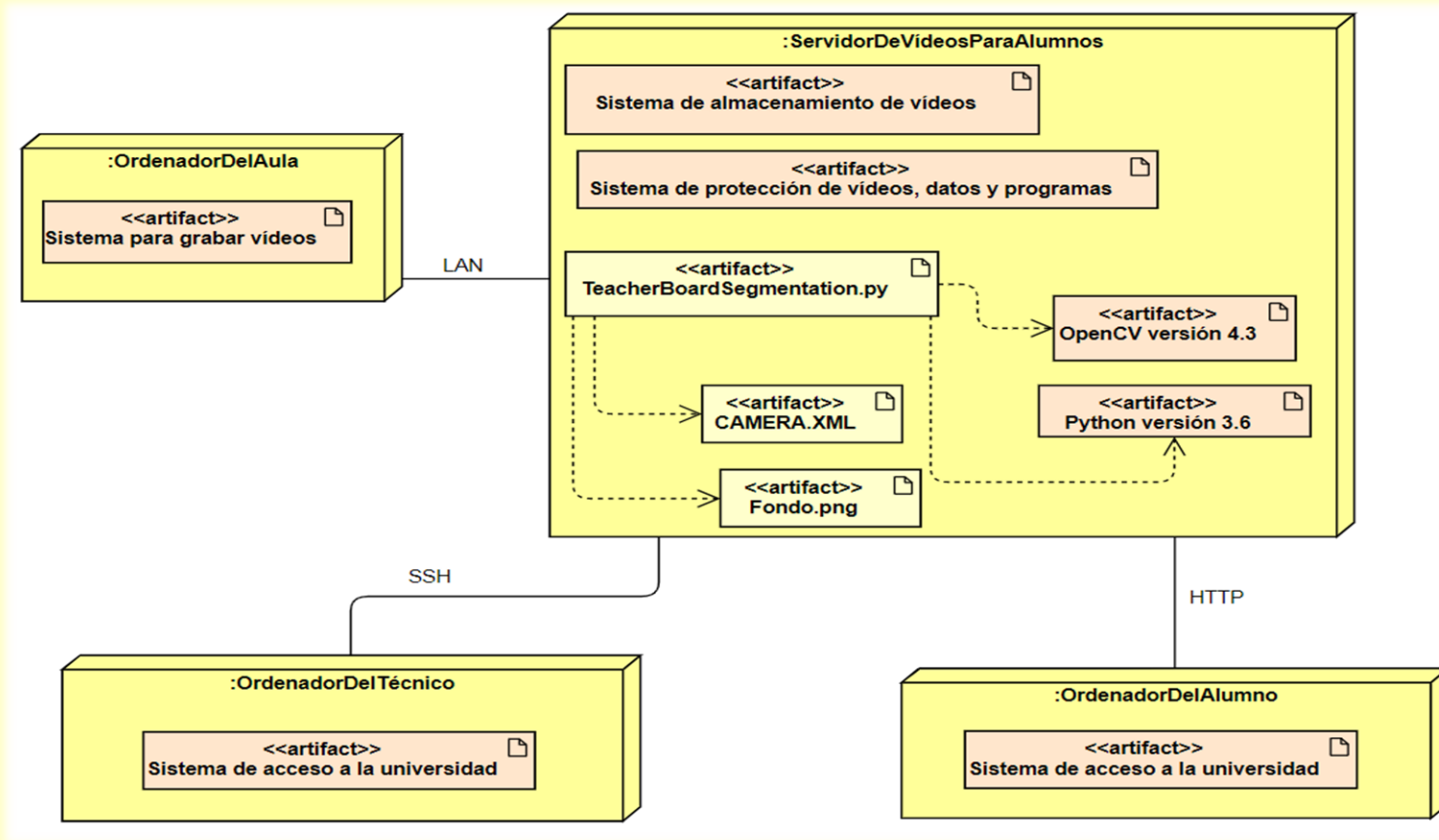


- Comprueba que están los **archivos de configuración** necesarios para el proceso.
- Segmenta al **profesor**.
- Segmenta la **pizarra** mejorando perspectiva y contraste.
- Genera los **vídeos de salida** tanto del profesor como de la pizarra.

4. DESCRIPCIÓN DEL ALGORITMO

Diagrama de despliegue:

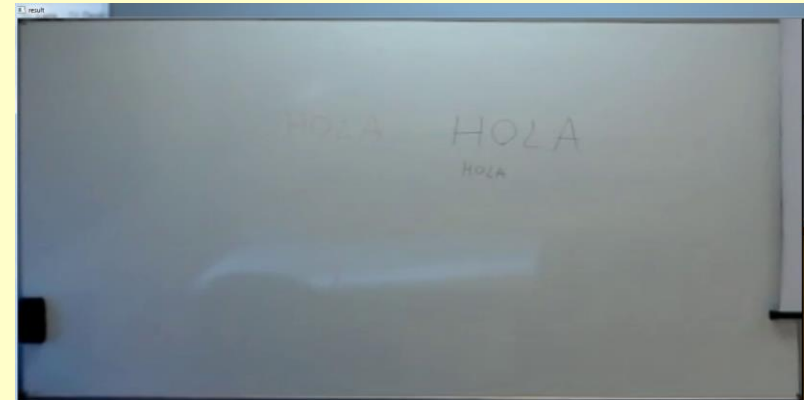
La arquitectura en tiempo de ejecución contará con **cuatro nodos** con sus respectivos **artefactos**.



5. PRUEBAS

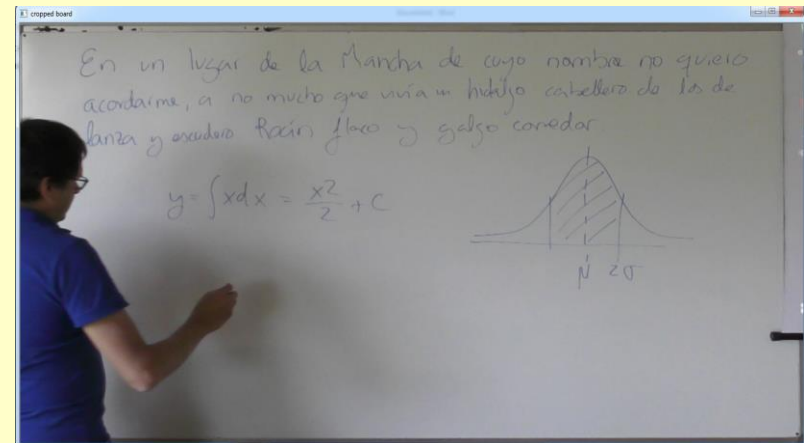
Se procesa un vídeo de 19 segundos de resolución **1280x720** con 454 frames de los cuales:

- 450 son **frames positivos** correspondientes al **profesor**.
- 4 son frames verdaderos negativos en los que no está el profesor.
- La **pizarra** está **localizada** en todos los frames.
- El **tiempo** de ejecución se corresponde con el del vídeo.



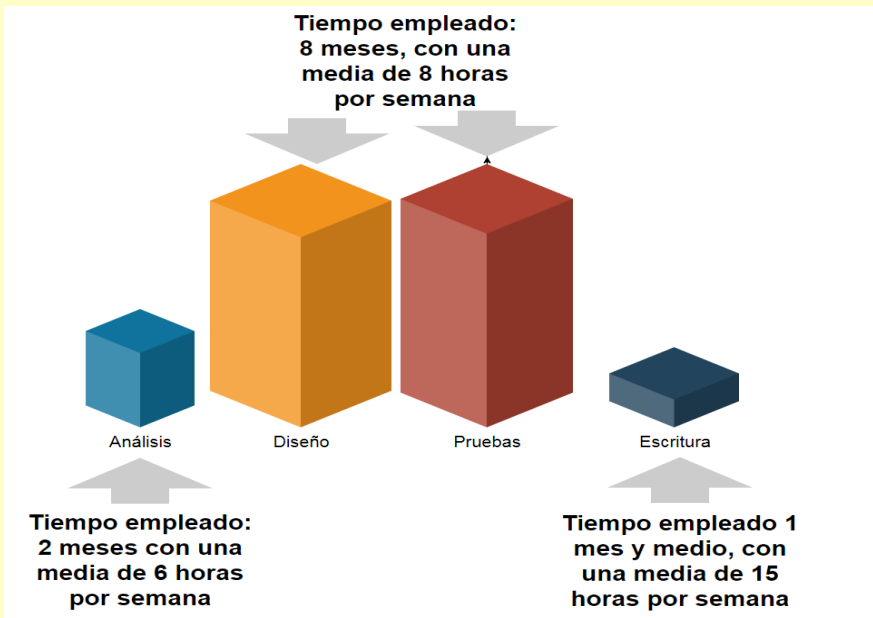
Se procesa un vídeo de 60 segundos en **4k** con 1525 frames de los cuales:

- 1524 son **frames verdaderos** positivos correspondientes al **profesor**.
- La **pizarra** está **localizada** en todos los frames.
- La ejecución tarda un **tiempo** de cuatro minutos.



6. MÉTRICAS

Teniendo en cuenta que el total de horas invertidas ha sido de **cuatrocientas horas** y suponiendo un coste de 20€/h, este proyecto que consta de **trescientas líneas** de código podría valorarse en 8000€.



7. CONCLUSIONES

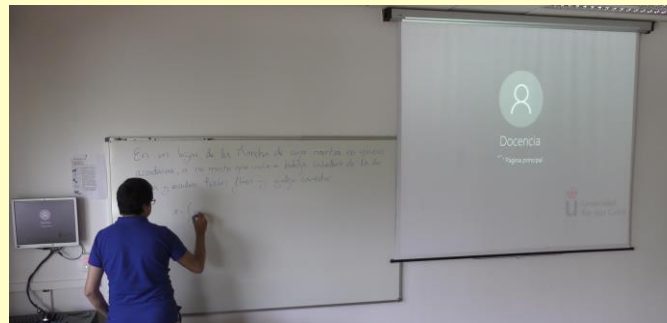
En este proyecto se ha construido un software que partiendo de un vídeo de una clase:

- **Detectada** la presencia del **profesor** y **segmentada** su imagen del resto de la escena.
- **Segmentada** la imagen de la **pizarra**, corrige su **perspectiva** y mejora su relación de contraste.

Cabe destacar sus posibles **mejoras futuras** como:

- Adaptación para **discapacitados** con subtítulos o sistema braille.
- Incorporación de un **script** para introducir archivos de configuración sin necesidad de realizar cambios en el código base.

Partiendo de un
vídeo de una clase



Vídeo de la
pizarra

Vídeo del
profesor



GRACIAS